

Entropy functions and determinant inequalities

Terence Chan, Dongning Guo, and Raymond Yeung

Abstract—In this paper, we show that the characterisation of all determinant inequalities for $n \times n$ positive definite matrices is equivalent to determining the smallest closed and convex cone containing all entropy functions induced by n scalar Gaussian random variables. We have obtained inner and outer bounds on the cone by using representable functions and entropic functions. In particular, these bounds are tight and explicit for $n \leq 3$, implying that determinant inequalities for 3×3 positive definite matrices are completely characterized by Shannon-type information inequalities.

Index Terms—Entropy, Gaussian distribution, rank functions

I. INTRODUCTION

Let n be a positive integer and denote the ground set by $\mathcal{N} = \{1, \dots, n\}$ throughout this paper. Suppose K is an $n \times n$ positive definite matrix. For any subset $\alpha \subseteq \mathcal{N}$, let K_α be the sub-matrix of K obtained by removing those rows and columns of K indexed by $\mathcal{N} \setminus \alpha$ and its determinant be denoted by $|K_\alpha|$. Note that when α is the empty set, we will simply define K_α as the scalar of value 1. There are many determinant inequalities in the existing literature that involve only the principle minors of the matrix. These include

- 1) Hadamard inequality

$$|K| \leq \prod_{i=1}^n |K_i| \quad (1)$$

- 2) Szasz inequality

$$\left(\prod_{\beta \subseteq \mathcal{N}: |\beta|=l} |K_\beta| \right)^{\frac{1}{\binom{k-1}{l-1}}} \geq \left(\prod_{\beta \subseteq \mathcal{N}: |\beta|=l+1} |K_\beta| \right)^{\frac{1}{\binom{k-1}{l}}} \quad (2)$$

for any $1 \leq l < k$.

As pointed out in [1], [2] and to be illustrated in Section II, many of such determinant inequalities (including the above two inequalities) can be proved via an information-theoretic approach. Despite that many determinant inequalities can be found in this approach, a complete characterisation of all determinant inequalities is still missing. In this paper, we aim to understand determinant inequalities by using the information inequality framework proposed in [3].

II. INFORMATION INEQUALITY FRAMEWORK

The framework proposed in [3] provides a geometric approach to understanding information inequalities.¹ Its idea will be illustrated shortly.

Definition 1 (Rank functions): A rank function over the ground set \mathcal{N} is a real-valued function defined on all subsets of \mathcal{N} . The rank function space over the ground set \mathcal{N} , denoted by \mathbb{R}^{2^n} , is the set of all rank functions over \mathcal{N} .

As usual, \mathbb{R}^{2^n} will be treated as a 2^n -dimensional Euclidean space, so that concepts such as metric and limits can be defined accordingly.

Definition 2 (Entropic functions): Let g be a rank function over \mathcal{N} . Then g is called *entropic* if there exists a set of discrete random variables $\{X_i, i \in \mathcal{N}\}$ such that $g(\alpha)$ is the Shannon entropy² $H(X_i, i \in \alpha)$, or $H(X_\alpha)$ for short, for all $\alpha \subseteq \mathcal{N}$.

On the other hand, if $\{X_i, i \in \mathcal{N}\}$ is a set of continuous scalar random variables such that $g(\alpha)$ is the differential entropy $h(X_\alpha)$ for all $\alpha \subseteq \mathcal{N}$, then g is called *s-entropic*.

Definition 3 (Entropic regions): Consider any nonempty finite ground set \mathcal{N} . Define the following “entropy regions”:

$$\Gamma_n^* = \{g \in \mathbb{R}^{2^n} : g \text{ is entropic}\} \quad (3)$$

$$\gamma_{s,n}^* = \{g \in \mathbb{R}^{2^n} : g \text{ is s-entropic}\}. \quad (4)$$

Understanding the above entropic regions is one of the most fundamental problems in information theory. It is equivalent to determining the set of all information inequalities [3].

In this paper, we will use the following notation. For any subset $\mathcal{S} \subseteq \mathbb{R}^{2^n}$, $\mathbb{W}(\mathcal{S})$ is defined as the set of all rank functions g^* such that $g^* = c \cdot g$ for some $c > 0$ and $g \in \mathcal{S}$. The closure of $\mathbb{W}(\mathcal{S})$ will be denoted by $\overline{\mathbb{W}}(\mathcal{S})$. Finally, the smallest closed and convex cone containing \mathcal{S} will be denoted by $\overline{\text{con}}(\mathcal{S})$. Clearly,

$$\mathcal{S} \subseteq \mathbb{W}(\mathcal{S}) \subseteq \overline{\mathbb{W}}(\mathcal{S}) \subseteq \overline{\text{con}}(\mathcal{S}). \quad (5)$$

Theorem 1 (Geometric framework [3]): A linear information inequality

$$\sum_{\alpha \subseteq \mathcal{N}} c_\alpha H(X_\alpha) \geq 0$$

is valid for all discrete random variables $\{X_1, \dots, X_n\}$ if and only if for all $g \in \Gamma_n^*$

$$\sum_{\alpha \subseteq \mathcal{N}} c_\alpha g(\alpha) \geq 0.$$

By Theorem 1, characterising the set of all valid information inequalities is thus equivalent to characterising the set Γ_n^* . Similar results can be obtained for the set $\gamma_{s,n}^*$. In the following, we will extend this geometric framework to study determinant inequalities.

¹See [4, Ch. 13-16] for a comprehensive treatment.

²All logarithms used in the paper is in the base 2.

Definition 4 (Log-determinant function): A rank function g over \mathcal{N} is called *log-determinant* if there exists an $n \times n$ positive definite matrix K such that

$$g(\alpha) = \log |K_\alpha| \quad (6)$$

for all $\alpha \subseteq \mathcal{N}$.

Let Ψ_n be the set of all log-determinant functions over \mathcal{N} . Then, we have the following theorem.

Theorem 2: Let $\{c_\alpha, \alpha \subseteq \mathcal{N}\}$ be any real numbers. The determinant inequality

$$\prod_{\alpha \subseteq \mathcal{N}} |K_\alpha|^{c_\alpha} \geq 1 \quad (7)$$

holds for all positive definite matrix K if and only if

$$\sum_{\alpha \subseteq \mathcal{N}} c_\alpha g(\alpha) \geq 0 \quad (8)$$

for all $g \in \overline{\text{con}}(\Psi_n)$.

Proof: By taking logarithm on both sides of the inequality, (7) is equivalent to that

$$\sum_{\alpha \subseteq \mathcal{N}} c_\alpha \log |K_\alpha| \geq 0 \quad (9)$$

for all positive definite matrix K . As (9) is a linear inequality, it is satisfied by all $g \in \Psi_n$ if and only if it is satisfied by all $g \in \overline{\text{con}}(\Psi_n)$. The theorem then follows. ■

In other words, the characterisation of the set of all determinant inequalities is equivalent to determining the set $\overline{\text{con}}(\Psi_n)$. In the rest of the paper, we will obtain inner and outer bounds on $\overline{\text{con}}(\Psi_n)$.

To achieve our goal, we will take an information theoretic approach [2]. The idea is very simple: Let $\{X_1, \dots, X_n\}$ be a set of scalar Gaussian random variables whose covariance matrix is equal to $(1/2\pi e)K$. Then the differential entropy of X_α is given by

$$h(X_\alpha) = \frac{1}{2} \log |K_\alpha|. \quad (10)$$

Definition 5 (Scalar Gaussian function): A function $g \in \mathbb{R}^{2^n}$ is called *s-Gaussian* if there exists scalar Gaussian variables $\{X_1, \dots, X_n\}$ where

$$g(\alpha) = h(X_\alpha) \quad (11)$$

for all $\alpha \subseteq \mathcal{N}$.

From (10), a rank function g is log-determinant if and only if $\frac{1}{2}g$ is *s-Gaussian*. Let $\Upsilon_{s,n}$ be the set of all *s-Gaussian* functions. Then

$$\overline{\text{con}}(\Psi_n) = \overline{\text{con}}(\Upsilon_{s,n}).$$

Consequently, we have the following theorem.

Theorem 3: The determinant inequality

$$\prod_{\alpha \subseteq \mathcal{N}} |K_\alpha|^{c_\alpha} \geq 1$$

holds for all positive definite matrix K if and only if

$$\sum_{\alpha \subseteq \mathcal{N}} c_\alpha h(X_\alpha) \geq 0$$

for all scalar Gaussian variables $\{X_1, \dots, X_n\}$.

In fact, the Hadamard inequality and Szasz inequality are respectively the counterparts of the following basic information inequalities³ [5]

$$\sum_{i=1}^n h(X_i) \geq h(X_1, \dots, X_n) \quad (12)$$

$$\frac{1}{\binom{k}{l}} \sum_{\beta \subseteq \mathcal{N}: |\beta|=l} \frac{h(Y_\beta)}{l} \geq \frac{1}{\binom{k}{l+1}} \sum_{\beta \subseteq \mathcal{N}: |\beta|=l+1} \frac{h(Y_\beta)}{l+1}. \quad (13)$$

In the following sections, we will obtain inner and outer bounds on the set $\overline{\text{con}}(\Upsilon_{s,n})$. The following corollaries of Theorem 2 show how these bounds can be used for proving or disproving a determinant inequality.

Corollary 1 (Proving an inequality): Suppose \mathcal{S} contains $\overline{\text{con}}(\Upsilon_{s,n})$ as a subset. The determinant inequality (7) holds for all positive definite matrix K if

$$\sum_{\alpha \subseteq \mathcal{N}} c_\alpha g(\alpha) \geq 0$$

for all $g \in \mathcal{S}$.

Therefore, any explicit outer bound on $\overline{\text{con}}(\Upsilon_{s,n})$ can lead to the discovery of new determinant inequalities. On the other hand, an inner bound on $\overline{\text{con}}(\Upsilon_{s,n})$ can be used for disproving a determinant inequality.

Corollary 2 (Disproving an inequality): Suppose $\mathcal{T} \subseteq \overline{\text{con}}(\Upsilon_{s,n})$. The determinant inequality (7) does not hold for all positive definite matrices if there exists $g \in \mathcal{T}$ such that

$$\sum_{\alpha \subseteq \mathcal{N}} c_\alpha g(\alpha) < 0.$$

III. AN INNER BOUND AND AN OUTER BOUND

As discussed earlier, log-determinant functions are essentially the same as *s-Gaussian* functions. Our objective is thus to characterise $\overline{\text{con}}(\Upsilon_{s,n})$, or at least to understand its basic properties. Since scalar Gaussian random variables are continuous scalar random variables, the next lemma follows immediately from the definition.

Lemma 1 (Outer bound):

$$\Upsilon_{s,n} \subseteq \gamma_{s,n}^*, \quad (14)$$

³Han's inequality was originally proved for discrete random variables. However, by using the same proving technique, it can also be proved to hold for all continuous random variables [1]. Alternative, its validity also follows from [6]: If a balanced information inequality (including Han's inequality) holds for all discrete random variables, then its "continuous counterpart" (i.e., the inequality by replacing discrete entropies with differential entropies) also holds for all continuous random variables.

and consequently,

$$\overline{\text{con}}(\Upsilon_{s,n}) \subseteq \overline{\text{con}}(\gamma_{s,n}^*). \quad (15)$$

It is well known that $\overline{\Gamma}_n^*$ (i.e., the closure of Γ_n^*) is a closed and convex cone [3]. It was established in [6] that

$$\overline{\text{con}}(\gamma_{s,n}^*) = \overline{\text{con}}(\overline{\Gamma}_n^*, \phi_1^n, \dots, \phi_n^n) \quad (16)$$

where

$$\phi_i^n(\alpha) = \begin{cases} -1 & \text{if } i \in \alpha \\ 0 & \text{otherwise.} \end{cases}$$

In the following, we prove an inner bound on $\overline{\text{con}}(\Upsilon_{s,n})$ by using representable functions.

Definition 6 (*s*-representable function): A rank function g over \mathcal{N} is called *s-representable* if there exists real-valued vectors (of the same length) $\{A_1, \dots, A_n\}$ such that for all $\alpha \subseteq \mathcal{N}$,

$$g(\alpha) = \dim \langle A_i, i \in \alpha \rangle.$$

In other words, $g(\alpha)$ is the maximum number of independent vectors in the set $\{A_i, i \in \alpha\}$.

Theorem 4 (Inner bound): If g is *s*-representable, then

$$g \in \overline{\text{con}}(\Upsilon_{s,n}).$$

Proof: Suppose the length of each row vector A_i is k . Let

$$\{W_1, \dots, W_k, V_1, \dots, V_n\}$$

be a set of independent standard Gaussian random variables. Therefore, its covariance matrix is the $(n+k) \times (n+k)$ identity matrix. Let $c > 0$. For each $i = 1, \dots, n$, define a real-valued continuous random variable as follows

$$X_i \triangleq \frac{1}{\sqrt{c}} A_i [W_1, \dots, W_k]^\top + V_i.$$

Let $\mathbf{X} = [X_1, \dots, X_n]^\top$. Then

$$\mathbf{X} = \frac{1}{\sqrt{c}} A [W_1, \dots, W_k]^\top + \mathbf{V}$$

where A is an $n \times k$ matrix whose i^{th} row is A_i and

$$\mathbf{V} = [V_1, \dots, V_n]^\top.$$

Since X_i is zero-mean,

$$\begin{aligned} \text{Cov}(\mathbf{X}) &= E[\mathbf{X}\mathbf{X}^\top] \\ &= \frac{1}{c} E[A[W_1, \dots, W_k]^\top [W_1, \dots, W_k] A^\top] + \mathbf{I} \\ &= \frac{1}{c} A A^\top + \mathbf{I}. \end{aligned}$$

Consequently,

$$\det(\text{Cov}(\mathbf{X})) = \det\left(\frac{1}{c} D + \mathbf{I}\right) \quad (17)$$

where D is the diagonal matrix obtained by using singular-value decomposition (SVD) over $A A^\top$. Let $d_1 \geq d_2 \geq \dots \geq d_n \geq 0$ be the diagonal entries of D and r be the rank of the

matrix $A A^\top$ (or equivalently, the rank of A). Hence, $d_i > 0$ if and only if $i \leq r$. Then

$$\det(\text{Cov}(\mathbf{X})) = \prod_{i=1}^r \left(\frac{d_i}{c} + 1 \right). \quad (18)$$

It is easy to see that

$$\lim_{c \rightarrow 0} \frac{h(X_1, \dots, X_n)}{\frac{1}{2} \log 1/c} = \lim_{c \rightarrow 0} \frac{\frac{1}{2} \log((2\pi e)^n \det(\text{Cov}(\mathbf{X})))}{\frac{1}{2} \log 1/c} \quad (19)$$

$$= \lim_{c \rightarrow 0} \frac{\log(\det(\text{Cov}(\mathbf{X})))}{\log 1/c} \quad (20)$$

$$= \lim_{c \rightarrow 0} \frac{\sum_{i=1}^r \log\left(\frac{d_i}{c} + 1\right)}{\log 1/c} \quad (21)$$

$$= r. \quad (22)$$

Similarly, for any $\alpha \subseteq \{1, \dots, n\}$, we can prove that

$$\lim_{c \rightarrow 0} \frac{h(X_\alpha)}{\frac{1}{2} \log 1/c} = \dim \langle A_i, i \in \alpha \rangle = g(\alpha).$$

Thus, $g \in \overline{\text{con}}(\Upsilon_{s,n})$ and the theorem is proved. \blacksquare

Lemma 2: Let $\{X_1, \dots, X_n\}$ be a set of scalar jointly continuous random variables with differential entropy function g . For any $c_1, \dots, c_n > 0$, define the set of random variables $\{Y_1, \dots, Y_n\}$ by

$$Y_i = X_i/c_i, \forall i \in \mathcal{N},$$

and let g^* be the differential entropy function of $\{Y_1, \dots, Y_n\}$. Then

$$g^*(\alpha) = g(\alpha) + \sum_{i \in \alpha} \log c_i \quad (23)$$

$$= g(\alpha) - \sum_{i \in \mathcal{N}} (\log c_i) \phi_i^n(\alpha) \quad (24)$$

for all $\alpha \subseteq \mathcal{N}$. Consequently, if g is *s*-Gaussian, then so is g^* .

Proof: Let f_{X_1, \dots, X_n} and f_{Y_1, \dots, Y_n} be respectively the probability density functions (pdfs) of $\{X_1, \dots, X_n\}$ and $\{Y_1, \dots, Y_n\}$. Then

$$\begin{aligned} f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) \\ = \left(\prod_{i=1}^n c_i \right) f_{X_1, \dots, X_n}(c_1 y_1, \dots, c_n y_n), \end{aligned} \quad (25)$$

and (23) can be directly verified. \blacksquare

Corollary 3:

$$\begin{aligned} \overline{\text{con}}(\Omega_{s,n}, \phi_1^n, \dots, \phi_n^n) &\subseteq \overline{\text{con}}(\Upsilon_{s,n}) \subseteq \overline{\text{con}}(\gamma_{s,n}^*) \\ &= \overline{\text{con}}(\overline{\Gamma}_n^*, \phi_1^n, \dots, \phi_n^n) \end{aligned} \quad (26)$$

where $\Omega_{s,n}$ is the set of all *s*-representable functions.

Proof: A direct consequence of Lemmas 1 and 2, Theorem 4 and (16). \blacksquare

Proposition 1 (Tightness of inner and outer bounds): For $n \leq 3$,

$$\begin{aligned} \overline{\text{con}}(\Omega_{s,n}, \phi_1^n, \dots, \phi_n^n) &= \overline{\text{con}}(\Upsilon_{s,n}) \\ &= \overline{\text{con}}(\gamma_{s,n}^*) = \overline{\text{con}}(\bar{\Gamma}_n^*, \phi_1^n, \dots, \phi_n^n). \end{aligned} \quad (27)$$

Proof: By Corollary 3, to prove the proposition, it suffices to prove that for $n \leq 3$,

$$\overline{\text{con}}(\bar{\Gamma}_n^*) \subseteq \overline{\text{con}}(\Omega_{s,n}). \quad (28)$$

In [16], the cone $\bar{\Gamma}_n^*$ (when $n \leq 3$) was explicitly determined by identifying the set of extreme vectors of the cone. It can be proved that all the extreme vectors are s -representable⁴ and hence is a subset of $\overline{\text{con}}(\Omega_{s,n})$. Consequently, (28) holds and the proposition follows. ■

Proposition 1 does not hold when $n \geq 4$. In fact, $\overline{\text{con}}(\Omega_{s,n}, \phi_1^n, \dots, \phi_n^n)$ is in general a proper subset of $\overline{\text{con}}(\Upsilon_{s,n})$ when $n \geq 4$. In [12], it was proved that all s -representable functions satisfy the Ingleton inequalities. It can also be directly verified that all the functions ϕ_i^n also satisfy the Ingleton inequalities. Therefore, all the functions in $\overline{\text{con}}(\Omega_{s,n}, \phi_1^n, \dots, \phi_n^n)$ also satisfy the Ingleton inequalities. However, in [10], it was proved that there exists $g \in \Upsilon_{s,n}$ for $n = 4$ that violates the the Ingleton inequality. Thus, $\overline{\text{con}}(\Omega_{s,n}, \phi_1^n, \dots, \phi_n^n)$ is indeed a proper subset of $\overline{\text{con}}(\Upsilon_{s,n})$.

IV. ANOTHER OUTER BOUND

By definition, the set $\overline{\text{con}}(\Psi_n)$ (which is the focus of our interest) is close under addition. However, this is not necessarily true for Ψ_n . In fact, $\bar{\mathbb{W}}(\Psi_n)$ is not necessarily equal to $\overline{\text{con}}(\Psi_n)$.

In the previous section, we showed that the set Ψ_n is essentially equivalent to the set of s -Gaussian functions, defined via sets of scalar Gaussian random variables. It turns out that, if we relax the constraint by allowing the Gaussian random variables to be vectors, instead of scalars, we will obtain an outer bound for Ψ_n and also $\overline{\text{con}}(\Psi_n)$.

Definition 7 (Vector Gaussian function): A function $g \in \mathbb{R}^{2^n}$ is called v -Gaussian if there exists n Gaussian random vectors $\{X_1, \dots, X_n\}$ such that

$$g(\alpha) = h(X_\alpha) \quad (29)$$

for all $\alpha \subseteq \mathcal{N}$.

Lemma 3: $\overline{\text{con}}(\Upsilon_{v,n}) = \bar{\mathbb{W}}(\Upsilon_{v,n})$.

Proof: It is clear from the definition that $\bar{\mathbb{W}}(\Upsilon_{v,n}) \subseteq \overline{\text{con}}(\Upsilon_{v,n})$. Now, consider positive integers k, ℓ_1, ℓ_2 and $g_1, g_2 \in \Upsilon_{v,n}$. It is easy to see that

$$\ell_1 g_1 + \ell_2 g_2 \in \Upsilon_{v,n}.$$

Hence,

$$\frac{\ell_1}{k} g_1 + \frac{\ell_2}{k} g_2 \in \mathbb{W}(\Upsilon_{v,n}).$$

Since k, ℓ_1, ℓ_2 are arbitrary positive integers, for any positive numbers $c_1, c_2 > 0$,

$$c_1 g_1 + c_2 g_2 \in \bar{\mathbb{W}}(\Upsilon_{v,n})$$

and the lemma follows. ■

Theorem 5 (Another outer bound):

$$\overline{\text{con}}(\Upsilon_{s,n}) \subseteq \bar{\mathbb{W}}(\Upsilon_{v,n}). \quad (30)$$

Proof: A direct consequence of that $\Upsilon_{s,n} \subseteq \Upsilon_{v,n}$ and Lemma 3. ■

So far, we have established two outer bounds (15) and (30) for $\overline{\text{con}}(\Upsilon_{s,n})$. In the following, we will prove that (30) is in fact a tighter one.

Definition 8: A rank function g is called v -entropic if there exists a set of random vectors $\{X_1, \dots, X_n\}$, not necessarily of the same length, such that

$$g(\alpha) = h(X_\alpha).$$

Also, let

$$\gamma_{v,n}^*(\mathcal{N}) = \{g \in \mathbb{R}^{2^n} : g \text{ is } v\text{-entropic}\}. \quad (31)$$

Clearly, $\bar{\mathbb{W}}(\Upsilon_{v,n}) = \overline{\text{con}}(\Upsilon_{v,n}) \subseteq \overline{\text{con}}(\gamma_{v,n}^*)$. Thus,

$$\overline{\text{con}}(\Upsilon_{s,n}) \subseteq \bar{\mathbb{W}}(\Upsilon_{v,n}) \subseteq \overline{\text{con}}(\gamma_{v,n}^*).$$

To show that (30) is tighter, it suffices to prove the following result.

Theorem 6: $\overline{\gamma_{v,n}^*} = \overline{\gamma_{s,n}^*} = \overline{\text{con}}(\bar{\Gamma}_n^*, \phi_1^n, \dots, \phi_n^n)$.

Theorem 6 basically states that replacing the real-valued random variables X_i in the vector \mathcal{X} by random vectors does not enlarge the closure of the space of differential entropy vectors. The discrete counterpart of this result is trivial, because as far as the probability masses and the entropy are concerned, a discrete random vector can be replaced by a scalar discrete random variable. However, in the continuous domain, it is not clear how a probability density function on \mathbb{R}^2 or more generally \mathbb{R}^m can be mapped to a pdf on \mathbb{R} without changing the entropies. In particular, there does not exist a continuous mapping from \mathbb{R}^2 to \mathbb{R} [9].

The proof of Theorem 6 exploits the relationship between the differential entropy of a continuous vector and the entropy of a discrete vector obtained through quantisation. Moreover, the entropy of the discrete random variable is equal to the differential entropy of a continuous random variable with piece-wise constant pdf. Given the n -tuple \mathbf{Z} whose entries are vectors, we “quantise” \mathbf{Z} by a discrete vector and then construct a continuous vector with n scalar entries whose entropy vector arbitrarily approximates that of \mathbf{Z} . Before we prove the theorem, we need several intermediate supporting results.

⁴In [16], the extreme vectors are proved to be representable with respect to a finite field. However, it can be verified easily that they are also s -representable with respect to the real field \mathbb{R} .

Lemma 4 (Closeness in addition): If g_1 and g_2 are v -entropic (or entropic) functions over \mathcal{N} , then their sum $g_1 + g_2$ is also v -entropic (or entropic).

Proof: Direct verification. ■

Proposition 2: If $g^* \in \gamma_{v,n}^*$, then for any $c > 0$, $c \cdot g^* \in \overline{\gamma_{v,n}^*}$.

Proof: Let $\mathbf{X} = (X_1, \dots, X_n)$ be a real-valued random vector with a probability density function. For any positive integer j , let $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(j)}$ be j independent replicas of \mathbf{X} (by a replica we mean a random object with identical distribution). Similarly, let $\mathbf{U} = (U_1, \dots, U_n)$ be a real-valued random vector such that U_1, \dots, U_n are mutually independent and each of them is uniformly distributed on the interval $[0, 1]$. Again, for any positive integer j , let $\mathbf{U}^{(1)}, \dots, \mathbf{U}^{(j)}$ be j independent replicas of \mathbf{U} . It is easy to see that the joint density function of $\mathbf{U}^{(1)}, \dots, \mathbf{U}^{(j)}$ is uniform on a hypercube with unit volume and hence has zero differential entropy.

Consider any $c > 0$. Let T be a binary random variable such that

$$\mathbb{P}\{T = 1\} = c/j \text{ and } \mathbb{P}\{T = 0\} = 1 - c/j$$

where j is a positive integer. Assume that T is independent of

$$(\mathbf{X}^{(1)}, \mathbf{U}^{(1)}, \dots, \mathbf{X}^{(j)}, \mathbf{U}^{(j)}).$$

Let $\mathbf{Z} = (Z_1, \dots, Z_n)$ where each Z_i is a random vector of length j such that for any $i = 1, \dots, n$,

$$Z_i = \begin{cases} (U_i^{(1)}, \dots, U_i^{(j)}) & \text{if } T = 0 \\ (X_i^{(1)}, \dots, X_i^{(j)}) & \text{otherwise.} \end{cases} \quad (32)$$

\mathbf{Z} is evidently continuous with a pdf, which is a mixture of two pdfs induced by that of \mathbf{X} and \mathbf{U} . For any $\alpha \subseteq \mathcal{N}$, we can directly verify that

$$h(Z_\alpha | T = 0) = h(U_\alpha^{(1)}, \dots, U_\alpha^{(j)}) \quad (33)$$

$$= 0 \quad (34)$$

and

$$h(Z_\alpha | T = 1) = h(X_\alpha^{(1)}, \dots, X_\alpha^{(j)}) \quad (35)$$

$$= jh(X_\alpha). \quad (36)$$

Consequently,

$$h(Z_\alpha | T) = ch(X_\alpha). \quad (37)$$

Hence,

$$ch(X_\alpha) = \lim_{j \rightarrow \infty} h(Z_\alpha | T) \quad (38)$$

$$\leq \lim_{j \rightarrow \infty} h(Z_\alpha) \quad (39)$$

$$\leq \lim_{j \rightarrow \infty} h(Z_\alpha | T) + h_b(c/j) \quad (40)$$

$$= ch(X_\alpha), \quad (41)$$

where $h_b(x)$ is the entropy of a binary random variable with probabilities x and $1 - x$. Thus, $\lim_{j \rightarrow \infty} h(Z_\alpha) = ch(X_\alpha)$. Let g^j and g^* be respectively the entropy function induced by

$\{Z_1, \dots, Z_n\}$ and $\{X_1, \dots, X_n\}$. Then g^j is v -entropic by definition and

$$\lim_{j \rightarrow \infty} g^j = c \cdot g^*.$$

Hence, $c \cdot g^* \in \overline{\gamma_{v,n}^*}$ for all $c > 0$ and our proposition follows. ■

Proposition 3: $\overline{\gamma_{v,n}^*}$ is a closed and convex cone.

Proof: For any $r \in \overline{\gamma_{v,n}^*}$, by definition, there exists a sequence of v -entropic functions $\{r^i\}_{i=1}^\infty$ such that

$$\lim_{i \rightarrow \infty} r^i = r.$$

Thus, for any $c > 0$,

$$\lim_{i \rightarrow \infty} c \cdot r^i = c \cdot r.$$

Then, by Proposition 2, $c \cdot r^i \in \overline{\gamma_{v,n}^*}$ and consequently, $c \cdot r \in \overline{\gamma_{v,n}^*}$.

Consider any $g_1^*, g_2^* \in \overline{\gamma_{v,n}^*}$, and $c_1, c_2 > 0$. Since

$$c_1 \cdot g_1^* \text{ and } c_2 \cdot g_2^* \in \overline{\gamma_{v,n}^*},$$

there exists sequences of v -entropic functions $\{r_1^i\}_{i=1}^\infty$ and $\{r_2^i\}_{i=1}^\infty$ such that

$$\lim_{i \rightarrow \infty} r_\ell^i = c_\ell \cdot g_\ell^*.$$

By Lemma 4, $r_1^i + r_2^i$ is also v -entropic. Thus,

$$c_1 \cdot g_1^* + c_2 \cdot g_2^* \in \overline{\gamma_{v,n}^*}.$$

The proposition is proved. ■

Definition 9 (m -Quantization): Given $m > 0$, let the m -quantization of any real number x be denoted as:

$$[x]_m = \frac{\lfloor mx \rfloor}{m} \quad (42)$$

where $\lfloor t \rfloor$ denotes the largest integer not exceeding t . Similarly, let the m -quantization of a real vector $\mathbf{x} = (x_1, \dots, x_n)$ be the element-wise m -quantization of the vector, denoted by $[\mathbf{x}]_m$, i.e.,

$$[\mathbf{x}]_m = ([x_1]_m, \dots, [x_n]_m). \quad (43)$$

Evidently, $[x]_m$ can only take values from the set

$$\left\{ 0, \pm \frac{1}{m}, \pm \frac{2}{m}, \dots \right\}. \quad (44)$$

Hence for every real-valued random variable X , $[X]_m$ is a discrete random variable taking value in the set (44). By definition,

$$\sum_{i \in \mathbb{Z}} \mathbb{P} \left\{ [X]_m = \frac{i}{m} \right\} = 1. \quad (45)$$

Proposition 4 (Renyi [7]): If X is a real-valued random vector of dimension n with a probability density function, then

$$\lim_{m \rightarrow \infty} H([X]_m) - n \log m = h(X). \quad (46)$$

Under the assumption that the pdf of a random variable X is Riemann-integrable, Proposition 4 is established in [8] by

treating $H([X]_m) - n \log m$ as the approximation of the Riemann integration of $-\int f_X(x) \log f_X(x) dx$. It is nontrivial to establish the result in general, where the pdf is not necessarily Riemann-integrable. An example of such a pdf can be defined by using the Smith-Volterra-Cantor set. Nonetheless (46) can be shown to hold using the Lebesgue convergence theorem along with some truncation arguments [7].

Lemma 5: Let $\{X_1, \dots, X_n\}$ be a set of discrete random variables such that its entropy function is g . For any positive numbers c_1, \dots, c_n , let g^* be defined as

$$g^*(\alpha) = g(\alpha) - \sum_{i \in \alpha} \log c_i.$$

Then g^* is s -entropic.

Proof: As X_i is discrete, we may assume without loss of generality that the sample space of X_i is the set of integers \mathbb{Z} . Let $p(x_1, \dots, x_n)$ be the probability mass function of $\{X_1, \dots, X_n\}$. Construct a set of continuous scalar random variables $\{Y_1, \dots, Y_n\}$ whose probability density function is defined as follows:

$$f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) \triangleq \left(\prod_{i=1}^n c_i \right) p(\lfloor c_1 y_1 \rfloor, \dots, \lfloor c_n y_n \rfloor).$$

It can then be directly verified that

$$h(Y_\alpha) = H(X_\alpha) - \sum_{i=1}^n \log c_i, \quad \forall \alpha \subseteq \mathcal{N}.$$

Consequently, g^* is s -entropic. \blacksquare

Proof of Theorem 6: Clearly, $\overline{\gamma_{s,n}^*} \subseteq \overline{\gamma_{v,n}^*}$. We will now prove that $\overline{\gamma_{v,n}^*} \subseteq \overline{\gamma_{s,n}^*}$. Let $\mathbf{Z} = (Z_1, \dots, Z_n)$ consist of n random vectors, where

$$Z_i = (Z_{i,1}, \dots, Z_{i,k_i}).$$

Let us define the m -quantization of Z_i , denoted as $[Z_i]_m$, be the element-wise m -quantization of Z_i , i.e., it consists of $[Z_{i,j}]_m$ for $j = 1, \dots, k_i$. By Proposition 4,

$$\lim_{m \rightarrow \infty} \left[H([Z_i]_m, i \in \alpha) - \left(\sum_{i \in \alpha} k_i \right) \log m \right] = h(Z_\alpha). \quad (47)$$

Let $g^*, r^m, g^m \in \mathbb{R}^{2^n}$ be such that

$$g^*(\alpha) = h(Z_\alpha) \quad (48)$$

$$r^m(\alpha) = H([Z_i]_m, i \in \alpha) \quad (49)$$

$$g^m(\alpha) = r^m(\alpha) - \left(\sum_{i \in \alpha} k_i \right) \log m. \quad (50)$$

By (47), $\lim_{m \rightarrow \infty} g^m = g^*$. Also, since $r^m \in \Gamma_n^*$, $g^m \in \gamma_{s,n}^*$ by Lemma 5. Consequently, $g^* \in \overline{\gamma_{s,n}^*}$. We have thus proved that $\overline{\gamma_{v,n}^*} \subseteq \overline{\gamma_{s,n}^*}$ and as a result, $\overline{\gamma_{v,n}^*} = \overline{\gamma_{s,n}^*}$. Finally, by Proposition 3, $\overline{\gamma_{v,n}^*}$ is a closed and convex cone and is equal to $\overline{\text{con}}(\gamma_{s,n}^*)$. Then by (16),

$$\overline{\gamma_{v,n}^*} = \overline{\text{con}}(\overline{\Gamma_n^*}, \phi_1^n, \dots, \phi_n^n). \quad (51)$$

The theorem is proved. \blacksquare

In Theorem 4, we have constructed an inner bound for $\overline{\text{con}}(\Upsilon_{s,n})$ by using s -representable functions. The same trick can also be used for constructing an inner bound for the set $\overline{\mathbb{W}}(\Upsilon_{v,n})$.

Definition 10: A rank function g over \mathcal{N} is called v -representable if for $i = 1, \dots, n$, there exists a set of real-valued vectors (of the same length) $\{A_{i,1}, \dots, A_{i,k_i}\}$ such that for all $\alpha \subseteq \mathcal{N}$,

$$g(\alpha) = \dim \langle A_{i,j}, i \in \alpha, j = 1, \dots, k_i \rangle.$$

The following theorem is a counterpart of Theorem 4. The proving technique is the same as before. We will omit the proof for brevity.

Theorem 7 (Inner bound on $\overline{\mathbb{W}}(\Upsilon_{v,n})$): Suppose that g is v -representable, then $g \in \overline{\mathbb{W}}(\Upsilon_{v,n})$.

Theorem 7 is of great interest. Characterising the set of v -representable functions have been a very important problem in linear algebra and information theory. It is also extremely difficult. For many years, it is only known that v -representable functions are polymatroidal and satisfies the Ingleton inequalities [11], [12]. The set of representable functions is only known when $n \leq 4$. However, there were some recent breakthrough in this areas. In [13], [14], many new subspace rank inequalities which are required to be satisfied by representable functions are discovered. In particular, via a computer-assisted mechanical approach, the set of all representable functions when $n \leq 5$ has been completely characterised. Interesting properties about the set of v -representable functions were also obtained [15]. Theorems 4 and 7 thus opens a new door to exploit results obtained about representable functions to characterise the set of Gaussian functions.

Corollary 4 (Inner bound on $\overline{\mathbb{W}}(\Upsilon_{v,n})$):

$$\overline{\text{con}}(\Omega_{v,n}, \phi_1^n, \dots, \phi_n^n) \subseteq \overline{\mathbb{W}}(\Upsilon_{v,n})$$

where $\Omega_{v,n}$ is the set of all v -representable functions.

Remark 1: While

$$\overline{\text{con}}(\Omega_{s,n}, \phi_1^n, \dots, \phi_n^n) \subseteq \overline{\text{con}}(\Upsilon_{s,n}),$$

it is still an open question whether

$$\overline{\text{con}}(\Omega_{v,n}, \phi_1^n, \dots, \phi_n^n) \subseteq \overline{\text{con}}(\Upsilon_{s,n})$$

or not.

We will end this section with a discussion of a related concept in a recent work [10]. Gaussian rank functions were studied in [10]. However, their definitions are slightly different from ours.

Definition 11 (Normalised joint entropy [10]): Let $\{X_1, \dots, X_n\}$ be a set of n jointly distributed vector valued Gaussian random variables such that each vector X_i is a vector of length T . Its *normalised Gaussian entropy function* g is a function in \mathbb{R}^{2^n} such that

$$g(\alpha) \triangleq \frac{1}{T} h(X_\alpha).$$

The only difference between Definitions 5 and 11 is the scalar multiplier $1/T$. Hence, a normalised Gaussian entropy function must be contained in the set $\mathbb{W}(\Upsilon_{v,n})$. In one sense, our proposed definition is slightly more general as we do not require all the random vectors X_i to have the same length. On the other hand, the “normalising factor” $1/T$ in Definition 11 can lead to some interesting results. For example, while we cannot prove that the closure of $\mathbb{W}(\Upsilon_{s,n})$ is closed and convex, [10] proved that the closure of the set of all normalised Gaussian entropy functions is indeed closed and convex.

Proposition 5: Let $\Upsilon_{N,n}^*$ ⁵ be the set of all normalised Gaussian entropy functions. Then

$$\overline{\text{con}}(\Upsilon_{N,n}^*) = \overline{\text{con}}(\Upsilon_{v,n}).$$

Proof: It can be directly verified from definitions that $\overline{\text{con}}(\Upsilon_{N,n}^*) \subseteq \overline{\text{con}}(\Upsilon_{v,n})$. Now, consider any $g \in \Upsilon_{v,n}$. Then by definition, there exists n Gaussian random vectors $\{X_1, \dots, X_n\}$ such that

$$g(\alpha) = h(X_\alpha) \quad (52)$$

for all $\alpha \subseteq \mathcal{N}$. Let ℓ_i be the length of the random vector X_i . Assume without loss of generality that $\ell_1 \geq \ell_i$ for all i .

Let $k = \sum_{i=1}^n (\ell_1 - \ell_i)$ and Y_1, \dots, Y_k be a set of scalar Gaussian random variables with identity covariance matrix and independent of $\{X_1, \dots, X_n\}$. For each $i = 1, \dots, n$, let $r_i = \sum_{j=1}^i (\ell_1 - \ell_i)$ and

$$Z_i = \begin{cases} X_i & \text{if } \ell_i = \ell_1 \\ (X_i, Y_{r_i+1}, \dots, Y_{r_{i+1}}) & \text{otherwise.} \end{cases}$$

Clearly, each Z_i is a Gaussian vector with the same length ℓ_1 . Let g^* be the normalised entropy function induced by $\{Z_1, \dots, Z_n\}$. It is easy to verify that $\ell_1 g^* = g$. Consequently, $\Upsilon_{v,n} \subseteq \overline{\text{con}}(\Upsilon_{N,n}^*)$ and the proposition thus follows. ■

Remark 2: Our Proposition 1 can also be derived from [10, Theorem 5], which proved that for any $g \in \Upsilon_{v,n}$ when $n = 3$, there exists a $\theta^* > 0$ such that for all $\theta \geq \theta^*$, $\frac{1}{\theta}g$ is vector Gaussian. However, their proof techniques are completely different.

V. CONCLUSION

In this paper, we took an information theoretic approach to study determinant inequalities for positive definite matrices. We showed that characterising all such inequalities for an $n \times n$ positive definite matrix is equivalent to characterising the set of all scalar Gaussian entropy functions for n random variables. While a complete and explicit characterisation of the set is still missing, we obtained inner and outer bounds respectively by means of linearly representable functions and vector Gaussian entropy functions.

It turns out that for $n \leq 3$, the set of all scalar Gaussian entropy functions is the same as the set of all differential entropy functions. The latter set is completely characterized

by Shannon-type information inequalities. Consequently, the aforementioned inner and outer bounds agree with each other. For $n \geq 4$, we showed the contrary, and the problem is seeming very difficult.

REFERENCES

- [1] T. Cover and J. Thomas, Elements of information theory, Wiley-Interscience, New York, NY, USA, 1991. ISBN 0-471-06259-6.
- [2] T. Cover, “Determinant inequalities via information theory,” *SIAM. J. Matrix Anal. & Appl.*, 9(3), pp.384-392, 1988.
- [3] R. W. Yeung, “A framework for linear information inequalities,” *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1924–1934, Nov 1997.
- [4] R. W. Yeung, *Information Theory and Network Coding*, Springer 2008.
- [5] T. S. Han, “Nonnegative entropy measures of multivariate symmetric correlations,” *Inform. Contr.*, 36: 133-156, 1978.
- [6] T. H. Chan, “Balanced information inequalities,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 3261 – 3267, 2003.
- [7] A. Rényi, *Probability Theory*. Budapest, Hungary: North Holland – Academiai Kiado, 1970.
- [8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, 2006.
- [9] K. Wiboon, “Bijections from \mathbb{R}^n to \mathbb{R}^m ,” available online at <https://www.math.lsu.edu/~kwiboon/talkpaper.pdf>, 2010.
- [10] B. Hassibi and S. Shadbakht, “The entropy region for three gaussian random variables,” in *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, July 2008, pp. 2634 –2638.
- [11] L. Guille, T. H. Chan, and A. Grant, “The minimal set of ingletton inequalities,” *accepted for publications in IEEE Trans. on Inform. Theory*, 2011.
- [12] A. W. Ingleton, “Representation of matroids.” London: Academic Press, 1971, pp. 149–167.
- [13] R. Dougherty, C. Freiling, and K. Zeger, “Linear rank inequalities on five or more variables,” *Arxiv preprint cs.IT/0910.0284v3*, 2009.
- [14] R. Kinser, “New inequalities for subspace arrangements,” *J. Combin. Theory Ser. A*, 2010.
- [15] T. Chan, A. Grant, and D. Pflüger, “Truncation technique for characterising linear polymatroids,” *accepted for publications in IEEE Trans. on Inform. Theory*, 2011.
- [16] Z. Zhang and R. W. Yeung, “On the characterization of entropy function via information inequalities,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 1440–1452, 1998.

⁵The subscript N is a mnemonic for the word “normalised”.